



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/593,275	06/13/2000	Upendra V. Chaudhari	YOR-2000-0168US1	7772
35195	7590	09/01/2006	EXAMINER	
FERENCE & ASSOCIATES 409 BROAD STREET PITTSBURGH, PA 15143			HAN, QI	
			ART UNIT	PAPER NUMBER
			2626	

DATE MAILED: 09/01/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No. 09/593,275	Applicant(s) CHAUDHARI ET AL.	
	Examiner Qi Han	Art Unit 2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 17 July 2006.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-27 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1-3,6-16 and 19-27 is/are rejected.
- 7) ☒ Claim(s) 4-5,17-18 is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
 2. ☐ Certified copies of the priority documents have been received in Application No. _____.
 3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- | | |
|--|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892) | 4) <input type="checkbox"/> Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____ |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | 5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152) |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
Paper No(s)/Mail Date _____ | 6) <input type="checkbox"/> Other: _____ |

DETAILED ACTION

1. The text of those sections of Title 35, U.S. Code not included in this action can be found in a prior Office action.

Response to Amendment

2. This communication is responsive to the applicant's amendment dated 07/17/2006. The Applicant(s) amended claims 1, 14 and 27 (see the amendment: pages 2-9).

The examiner withdraws the claim rejection under 35 USC 112 2nd, because the applicant amended the claim.

Response to Arguments

3. Applicant's arguments filed on 07/17/2006 with respect to the rejection of claims 1-27 under 35 USC 102 and/or 103, fully considered but they are not persuasive.

In response to applicant's argument regarding rejection of claims 1-3, 6-2, 14-16 19-25 and 27 under 35 USC 102/103 that there is no suggestion or motivation to combine the references (see the amendment: pages 11, paragraph 4 to page 13, paragraph 3), examiner recognizes that obviousness can only be established by combining or modifying the teachings of the prior art to produce the claimed invention where there is some teaching, suggestion, or motivation to do so found either in the references themselves or in the knowledge generally available to one of ordinary skill in the art. See *In re Fine*, 837 F.2d 1071, 5 USPQ2d 1596 (Fed. Cir. 1988) and *In re Jones*, 958 F.2d 347, 21 USPQ2d 1941 (Fed. Cir. 1992). In this case, the

Art Unit: 2626

obviousness is based on the common knowledge in the art and/or the prior teachings. It is noted that providing multiple phonetic detail levels as a set of speech models for increasing efficiency and/or quality of speech or speaker recognition is a common knowledge in the art, because it provides more detailed information of a series of speech units for the recognition, so as to get better recognition result. It is also noted that both references (Goldenthal and Newman) work in the same field (speaker recognition) of endeavor and provides detailed acoustic features and models to determine the likelihood scores for the recognition (to solve the same problem). Further, Goldenthal teaches using ‘a pattern classification and recognition methodology’ and ‘to accurately represent and account for the dynamic behavior of the acoustic attributes’ (col. 4, lines 8-32) (which is incorporated by reference of US 5,625,749 that includes multiple detailed levels of speech processing: see US 5,625,749: Fig. 9 and col. 17, lines 1-12), and Newman teaches ‘identifying speaker’ ‘using speech recognition’ and using ‘the speech model that most closely matches the sample of speech for the unidentified speaker’ (col. 2, line 66 to col. 3, line 17); wherein both references provide the suggestion and/or motivation of using more detailed information for obtaining better speaker recognition result.

In response to applicant's argument regarding claim 1, that the prior art references (Goldenthal and Newman) “do not teach or suggest all of the claim limitations”, Newman “does not involve **hierarchical resolution**” and “does not compare a speech signal to a speech model at **plurality of level of phonetic detail of varying resolution**”, so that “the examiner’s rejection is therefore improper” (see the amendment: page 13, paragraph 4 to page 15, paragraph 2), the examiner respectfully disagrees with the applicant and has a different view of the prior art teachings and the claim interpretations. It should be pointed out that the argued and claimed

Art Unit: 2626

terms “**hierarchical resolution**” and “plurality of level of phonetic detail of **varying resolution**” are not specifically defined or described in the original specification, so that, as best understood, the terms are interpreted as “**levels of phonetic detail**”, in light of specification (see the closest disclosure in the specification: page 4, lines 9-10 and the original claim 1).

Further, it is noted that, as rejected in the office action, the combined references disclose all claimed limitation, based on the interpretation of above argued and claimed terms, because Newman discloses that 'each word 700 (Fig. 7) is represented by a set of phonemes 705 that represent the phonetic spelling of the word', and 'each phoneme is represented by three sets of model parameters 710 that correspond to the three nodes of the phoneme' (the model hierarchically resolved) (column 6, lines 29-34), and 'comparing each frame of the sequence of frames to model parameters from retrieved model (target model) for the phoneme node' (col. 6, line 66 to col. 7, line 2), which suggests that the system includes multiple levels of phonetic detail and the corresponding processing for each level (the model hierarchically resolved, also see Figs. 3 and 7), as claimed.

Furthermore, Newman discloses using 'a dynamic programming techniques to identify the series of words', which can also be read on the claimed and argued limitation of “**level of phonetic detail**” (also corresponding to “the target speaker model interpreted as hierarchical resolution” and “plurality of level of phonetic detail of varying resolution”), based on broadest reasonable interpretation of the claimed limitation, because dynamic programming for speech/speaker processing necessarily involves processing at least two levels of phonetic detail, which is well known in the art.

In addition, as mentioned above, Goldenthal teaches using ‘a pattern classification and recognition methodology’ (col. 4, lines 8-32), which is incorporated by reference of US 5,625,749 that includes multiple levels of detail for speech processing (see US 5,625,749: Fig. 9 and col. 17, lines 1-12). This means that Goldenthal’s system has capability of using multiple levels of phonetic detail for speaker processing incorporated by the reference, which further supports the claim rejection for the argued limitation. Therefore, the claimed limitation cannot be distinguishable from the prior art teachings and cannot overcome the obviousness of the claim rejection based on the combined references.

Regarding other claims, the response is based on the same reason describe for claim 1, because the related arguments are based the same issue discussed above.

For above reason, it is believed that the combined claim rejection is proper and the applicant’s arguments are not persuasive. The rejection is sustained.

Claim Rejections - 35 USC § 103

4. Claims 1-3, 6-13, 14-16 and 19-27 are rejected under 35 U.S.C. 103(a) as being unpatentable over Goldenthal et al. (US 6,205,424) hereinafter referenced as Goldenthal, in view of Newman et al. (US 5,946,654) hereinafter referenced as Newman.

Regarding **claim 1**, Goldenthal discloses two-staged cohort selection for speaker verification system (title), comprising:

“providing a model corresponding to a target speaker, the model being resolved [hierarchically] into at least one frame”, (column 3, line 64 to column 4, line 29, ‘the frames...processed by a model generator to produce sets of acoustic models which characterize

Art Unit: 2626

the original speech signals', 'one set of acoustic models for every identified speaker (target speaker) desiring enrollment');

“receiving an identity claim”, (column 1, lines 47-49, 'the claimed identity of an individual can be verified by having the individual utter a prompted sequence of words or spontaneous speech during a testing session');

“ascertaining whether the identity claim corresponds to the target speaker model”, (column 1, lines 56-57, 'if the score exceed a predetermined threshold its presumed that the individual is who he or she claims to be');

“said ascertaining step comprising the steps of: determining, for each frame [and each level] of phonetic detail of the target speaker model, a likelihood value; and resolving the at least one likelihood value to obtain a likelihood score”, (column 1, lines 50-57, 'these validation or testing speech signals are analyzed and compared with the pre-stored observation models corresponding to the "claimed" identity to determine scores', 'the scores can be expressed as log likelihood scores: $\text{score} = \log p(O/I)$, where p represents the likelihood that the observed frames O were produced by the individual I).

But, Goldenthal does not expressly disclose the model being resolved “hierarchically”, the frame(s) “comprising a plurality of levels of phonetic detail of varying resolution”, and determining a likelihood value for “each level” of the phonetic detail of target speaker model. However, these features are well known in the art as evidenced by Newman who, in the same field of endeavor, discloses speaker identification using unsupervised speech models (title), comprising that that 'each word 700 (Fig. 7) is represented by a set of phonemes 705 that represent the phonetic spelling of the word', and 'each phoneme is represented by three sets of

Art Unit: 2626

model parameters 710 that correspond to the three nodes of the phoneme' (the model hierarchically resolved) (column 6, lines 29-34), and 'comparing each frame of the sequence of frames to model parameters from retrieved model for the phoneme node' (col. 6, line 66 to col. 7, line 2), which suggests that the system includes multiple levels of phonetic detail and the corresponding processing for each level (the model hierarchically resolved), as claimed.

Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Goldenthal by specifically providing multiple phonetic detail levels and the corresponding processing for the model hierarchically resolved, as taught by Newman, for the purpose (motivation) of increasing efficiency and quality of a recognition system.

Regarding **claim 2** (depending on claim 1), Goldenthal in view of Newman further discloses "for each frame and each level of phonetic detail likelihood value is a maximum likelihood value" (Goldenthal: column 1, lines 53-54, 'the a log likelihood score'; column 2, lines 21-31, the log likelihood 'function f can be statistical ... maximum').

Regarding **claim 3** (depending on claim 2), Goldenthal in view of Newman further discloses "said step of resolving the at least one likelihood value comprises averaging the at least one likelihood value", (Goldenthal: column 1, lines 53-54, 'the a log likelihood score'; column 2, lines 21-31, the log likelihood 'function f can be statistical ... average').

Regarding **claim 6** (depending on claim 1), Goldenthal in view of Newman further discloses "the at least one level of phonetic detail comprises at least one of the following: a global level; a phonemic level and a subphonemic level", (Goldenthal: column 4, lines 8-29, 'a segment based speech approach to speech processing' and 'that designated segment can be units of speech, for example, phones, or transition from one phone to another').

Regarding **claim 7** (depending on claim 6), as stated above (see claim 1), Goldenthal in view of Newman discloses “the at least one level of phonetic detail comprises all of the following three levels: a global level; a phonemic level and a sub-phonemic level” (Newman: column 6, lines 29-34, 'each word 700 (Fig. 7) is represented by a set of phonemes 705 that represent the phonetic spelling of the word, and each phoneme is represented by three sets of model parameters 710 that correspond to the three nodes of the phoneme', which reads on the claim).

Regarding **claim 8** (depending on claim 7), Goldenthal fails to expressly disclose “providing labeling information for each frame.” However, the feature is well known in the art as evidenced by Newman who further discloses the labeling information in Figs 5-6 and 8. Therefore, it would have been obvious to one of ordinary skill in the art at the time the invention was made to modify Goldenthal by specifically providing labeling information for each frame, as taught by Newman, for the purpose of increasing efficiency of a recognition system.

Regarding **claim 9** (depending on claim 1), Goldenthal in view of Newman further discloses “accepting or rejecting the identity claim”, (Goldenthal: column 1, lines 50-57, 'if the scores exceed a predetermined threshold, it is presumed that the individual is who he or she claims to be'; Newman: column 2, line 44, 'Bayesian adaptation approach'; which necessarily includes accepting or rejecting the identity claim).

Regarding **claim 10** (depending on claim 1), as stated above, Goldenthal in view of Newman discloses “comparing a quantity based on the likelihood score to a predetermined threshold value”, (Goldenthal: column 1, lines 50-57, 'if the scores exceed a predetermined threshold, it is presumed that the individual is who he or she claims to be').

Regarding **claim 11**(depending on claim 10), Goldenthal in view of Newman further discloses “the steps of providing at least one model corresponding to at least one background speaker; and determining the quantity based on the likelihood score via employing the at least one background speaker model”, (Goldenthal: column 4, lines 49-58, 'a plurality of sets of "cohort" models (CM) 170 (Fig. 1) which characterize the speech signals of each identified speaker, are selected from the available sets of acoustic models of the other speakers', 'the selection can be made according to predetermined selection criteria, for example, the models which best characterize the speech of the identified speaker, or the models whose characterization fits some predetermined probability density function', which suggests that the combined system has capability of implementing the functionality as claimed).

Regarding **claim 12** (depending on claim 11), Goldenthal in view of Newman further discloses “said step of determining the quantity based on the likelihood comprises determining a log-likelihood ratio based on the likelihood score”, (Goldenthal: column 2, lines 21-28, 'that during testing, the score obtained from the models of the speaker whose identity is claimed is compared with all of the scores derived from the small set of cohort models to produce a set of score differences, and the differences are then used as a normalized score = $\log p(O/I) - f[\log p(O/(C_k(I))]$, where $\log p(O/(C_k(I))$ are the scores for the k cohorts linked to the claimed individual').

Regarding **claim 13** (depending on claim 12), Goldenthal in view of Newman further discloses “the log-likelihood ratio is determined by” the claimed equation, (Goldenthal: column 2, lines 21-31, 'a function f can **combine** all of the cohort scores during the normalization' and 'the function can be statistical in nature, for example, ...**average**...', which can be read on the

Art Unit: 2626

claim, based on the applicant's statement regarding the variables in the claimed equation in the amendment filed on 10/25/2005, on page 13, paragraph 2).

Regarding **claims 14-16 and 19-26**, they recite an apparatus. The rejection is based on the same reason described for claims 1-3 and 6-13, respectively, because claims 14-16 and 19-26 recite same or similar limitation(s) as claims 1-3 and 6-13, respectively.

Regarding **claim 27**, it discloses a program storage device readable by machine, which corresponds to the method of claim 1. The rejection is based on the same reason described for claim 1 because the claim recites same or similar limitation(s) as claim 1.

Allowable Subject Matter

5. Claims 4-5 and 17-18 are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

The following is an examiner's statement of reasons for the allowable subject matter:

Regarding claim 4, the prior art of record fail to specifically disclose or fairly suggest a way to determine the likelihood value through a particular equation, as described in the claim, which calculates the likelihood score by using multiple levels of phonetic detail of the speaker model, each level may have multiple processing units, wherein the multiple levels (L) is interpreted as more than one level in most of processing situation.

Regarding claim 5, it is dependent claim of the claim 4 and includes all features of its parent claims).

Art Unit: 2626

The prior art of record provided numerous teachings of alternative types of speaker recognition, identification and verification. However, the features as presented above are not anticipated by, nor made obvious over the prior art of the record.

6. Any comments considered necessary by applicant must be submitted no later than the payment of the issue fee and, to avoid processing delays, should preferably accompany the issue fee. Such submissions should be clearly labeled "Comments on Statement of Reasons for Allowance."

Conclusion

7. Please address mail to be delivered by the United States Postal Service (USPS) as follows:

Mail Stop ____
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

or faxed to: 571-273-8300, (for formal communications intended for entry)

Or: 571-273-8300, (for informal or draft communications, and please label "PROPOSED" or "DRAFT")

If no Mail Stop is indicated below, the line beginning Mail Stop should be omitted from the address.

Effective January 14, 2005, except correspondence for Maintenance Fee payments, Deposit Account Replenishments (see 1.25(c)(4)), and Licensing and Review (see 37 CFR 5.1(c) and 5.2(c)), please address correspondence to be delivered by other delivery services (Federal Express (Fed Ex), UPS, DHL, Laser, Action, Purolater, etc.) as follows:

U.S. Patent and Trademark Office
Customer Window, Mail Stop ____
Randolph Building
Alexandria, VA 22314

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Qi Han whose telephone numbers is (571) 272-7604. The examiner can normally be reached on Monday through Thursday from 9:00 a.m. to 7:00 p.m. If

Art Unit: 2626

attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil, can be reached on (571) 272-7602.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Inquiries regarding the status of submissions relating to an application or questions on the Private PAIR system should be directed to the Electronic Business Center (EBC) at 866-217-9197 (toll-free) or 703-305-3028 between the hours of 6 a.m. and midnight Monday through Friday EST, or by e-mail at: ebc@uspto.gov. For general information about the PAIR system, see <http://pair-direct.uspto.gov>.

QH/qh

August 30, 2006



RICHEMOND DORVIL
SUPERVISORY PATENT EXAMINER